

EXHIBIT 37
[FILED UNDER SEAL]

Tracking, Profiling, and Ad Targeting in the Alexa Echo Smart Speaker Ecosystem

Umar Iqbal
Washington University in St. Louis

Pouneh Nikkhah Bahrami
University of California, Davis

Rahmadi Trimananda
University of California, Irvine

Hao Cui
University of California, Irvine

Alexander Gamero-Garrido
University of California, Davis

Daniel J. Dubois
Northeastern University

David Choffnes
Northeastern University

Athina Markopoulou
University of California, Irvine

Franziska Roesner
University of Washington

Zubair Shafiq
University of California, Davis

ABSTRACT

Smart speakers collect voice commands, which can be used to infer sensitive information about users. Given the potential for privacy harms, there is a need for greater transparency and control over the data collected, used, and shared by smart speaker platforms as well as third party skills supported on them. To bridge this gap, we build a framework to measure data collection, usage, and sharing by the smart speaker platforms. We apply our framework to the Amazon smart speaker ecosystem. Our results show that Amazon and third parties, including advertising and tracking services that are unique to the smart speaker ecosystem, collect smart speaker interaction data. We also find that Amazon processes smart speaker interaction data to infer user interests and uses those inferences to serve targeted ads to users. Smart speaker interaction also leads to ad targeting and as much as 30× higher bids in ad auctions, from third party advertisers. Finally, we find that Amazon's and third party skills' data practices are often not clearly disclosed in their policy documents.

CCS CONCEPTS

• Security and privacy → Human and societal aspects of security and privacy; • Social and professional topics → Privacy policies; Corporate surveillance;

KEYWORDS

Alexa, Echo, Amazon, Smart speaker, Data collection, Targeted advertising, Corporate surveillance

ACM Reference Format:

Umar Iqbal, Pouneh Nikkhah Bahrami, Rahmadi Trimananda, Hao Cui, Alexander Gamero-Garrido, Daniel J. Dubois, David Choffnes, Athina Markopoulou, Franziska Roesner, and Zubair Shafiq. 2023. Tracking, Profiling, and Ad Targeting in the Alexa Echo Smart Speaker Ecosystem. In *Proceedings of the 2023 ACM Internet Measurement Conference (IMC '23)*, October 24–26, 2023, Montreal, QC, Canada. ACM, New York, NY, USA, 15 pages. <https://doi.org/10.1145/3618257.3624803>



This work is licensed under a Creative Commons Attribution International 4.0 License.

IMC '23, October 24–26, 2023, Montreal, QC, Canada
© 2023 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0382-9/23/10.
<https://doi.org/10.1145/3618257.3624803>

2023, Montreal, QC, Canada. ACM, New York, NY, USA, 15 pages. <https://doi.org/10.1145/3618257.3624803>

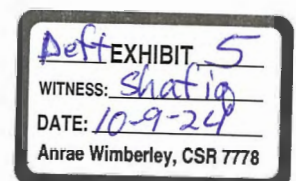
1 INTRODUCTION

Smart speakers are becoming increasingly prevalent because of their convenience [71]. However, with that convenience come privacy risks, since the data collected by smart speakers (i.e., voice recordings, their transcripts, and interaction metadata) can reveal, or can be used to infer, sensitive information about the users. For example, smart speaker vendors or third parties may infer users' sensitive physical (e.g., age, health) and psychological (e.g., mood, confidence) traits from their voice recordings [68]. Similarly, the set of questions and commands issued to a smart speaker may reveal sensitive information about users' states of mind, interests, and concerns. Despite the significant potential for privacy harms, users have little-to-no visibility into what information is captured by smart speakers, how it is shared with other parties, or how it is used by such parties.

Prior work provides ample evidence to support the need for greater transparency into smart speaker data collection, usage, and sharing. For instance, smart speaker platforms have been previously found to host malicious third party apps [22, 80], to mis-activate [26], and record users' private conversations without their knowledge [32, 34], and to share users' conversations with strangers [67]. There is a clear need to audit how smart speaker ecosystems handle data from their users' interactions.

To facilitate such independent, repeatable audits, we need an approach that can work on unmodified, off-the-shelf devices, without relying on cooperation by the smart speaker manufacturer. Conducting such audits requires addressing two key open challenges. First, commercially available smart speakers are closed-box devices without open interfaces that would allow independent researchers to expose what data is collected. Second, once data gathered from a smart speaker is sent over the Internet, there is no way to track how the data is further shared and used.

In this paper, we address these challenges by building an auditing framework that measures the collection, usage, and sharing of smart speaker interaction data. Our key insight is that data collection, usage, and sharing can be: (i) directly observed by intercepting the network traffic on the router without modifying smart speakers and



(ii) *indirectly inferred through its usage in targeted advertisements*. To operationalize this insight, we first expose data to smart speakers, observe the endpoints contacted during that exposure to capture the online services that collect data, and then analyze the targeted advertisements to infer the usage and sharing of the exposed data. We apply our framework to the Alexa Echo smart speaker ecosystem, which is the largest ecosystem, with 46 million devices in the US [72] and more than 200K third party applications [30].

To expose data, we simulate several treatment and control *personas* with different smart speaker usage profiles. Each treatment persona is simulated by installing and interacting with *skills* (the term for apps in the Alexa Echo ecosystem) from different categories on separate Alexa Echos, according to personas (e.g., a “fashion” persona is configured by installing and interacting with skills from the fashion category). By contrast, in the control persona, we do not install and interact with skills on Alexa Echo, and consequently do not expose any data.

To measure data collection, we build a custom Raspberry Pi (RPi) router [63] that allows us to capture the network endpoints contacted, by unmodified off-the-shelf Alexa Echos, during the skill installation and interaction process. To infer data usage and sharing, we look for statistically significant differences in the online targeted advertising between the treatment and control personas [21, 24, 60]. We measure targeting across two modalities: ad auction bid values and ad content served to the personas. By comparing ad auction bid values and ad content across treatment (i.e., when data is exposed) and control (i.e., when data is not exposed) personas, we can identify when smart speaker interactions are likely the cause of ad targeting, and thus infer that the data was used and/or shared for that purpose.

We summarize our key contributions as follows:

- (1) We develop a novel framework, that relies on direct and indirect measurements to understand the data collection, usage, and sharing in the Alexa Echo smart speaker ecosystem, without relying on cooperation from the smart speaker manufacturer.
- (2) We find that Alexa Echo interaction data is collected by both Amazon and third parties, including advertising and tracking services. As many as 41 advertisers sync their cookies (i.e., share data) with Amazon, and further with 247 other third parties, including advertising services.
- (3) We find evidence that Amazon processes Echo interactions to infer user interests, which was not clearly stated in Amazon’s policies before our research and public disclosure. Our measurements also indicate that inferred interests are used to serve targeted ads to users on the web. Advertisers bid as much as 30× higher on some personas.
- (4) Third party skills often do not clearly disclose their data collection practices in their privacy policies. For example, only 10 (2.2%) skills clearly disclose the endpoints that collect data and 68.61% (129/188) of skills do not even mention Alexa or Amazon in their privacy policies.

Paper organization. The rest of this paper is organized as follows: Section 2 provides background and motivation behind our research. Section 3 presents our proposed auditing framework. Section 4 presents the results of network traffic analysis to measure data collection. Section 5 presents the ad targeting analysis to infer

data usage and sharing. Section 6 analyzes the consistency of privacy policies, and other disclosures, with observed data collection, usage, and sharing practices. Section 7 provides discussion and Section 8 concludes the paper.

2 BACKGROUND & MOTIVATION

2.1 Alexa & Echo

In this paper, we study Alexa Echo smart speaker ecosystem, the most widely used ecosystem with more than 46 million devices in the US [72]. Echos are Alexa-powered smart speakers from Amazon. Alexa is a voice assistant that responds to user requests conveyed through voice input. Although Alexa can respond to a wide variety of general-purpose requests, it is not well-suited for specialized tasks, e.g., ordering a pizza from a particular restaurant. Thus, to augment Alexa, Amazon allows third party services to build and publish applications called *skills* on the Alexa marketplace. Alexa marketplace hosts more than 200K third party skills [30].

2.2 Privacy issues

The inclusion of third party skills poses a privacy risk to the users of Alexa Echos. Accordingly, Amazon imposes a set of platform policies to mitigate potential privacy risks of third party skills. Amazon restricts skills from collecting sensitive information, e.g., social security and bank account numbers [9, 10], and requires user permission to allow access to personal information, e.g., email, phone, location [15]. To enforce the aforementioned policies, Amazon has a skill certification process that aims to filter malicious skills before they can be published on the marketplace [8]. However, prior research has shown that policy-violating skills can get certified [22] and thousands of skills on the Alexa marketplace violate platform policies [80].

Smart speaker platforms, such as the Alexa Echo, also store voice recordings, their transcripts, and the information about the resulting action (metadata) generated from the user’s voice commands. This data can be used to infer sensitive information about the user. For example, user commands in their raw form, i.e., voice recording, can be used to infer several physical (e.g., age, health) and psychological characteristics (e.g., mood, confidence) of the user [68]. User commands in their processed form, i.e., transcripts of voice recordings, can expose sensitive information (e.g., private conversations) about the user. Even the metadata that results in execution of a user command (e.g., a user querying a third party health-related skill) can leak sensitive information about the user. Amazon aims to limit some of these privacy issues through its platform design choices [6]. For example, to avoid recording private conversations, user commands are only recorded when a user utters the *wake word*, e.g., Alexa (though prior research has also shown that smart speakers often *misactivate* and unintentionally record conversations [26]). Additionally, Amazon currently does not directly use users’ voice recordings to serve targeted advertisements [52], despite patenting that idea [47]. Further, only keywords from transcripts of user commands are shared with third-party skills, instead of the raw audio [17].

The limiting of voice recordings and their full transcripts reduce the privacy risks posed to users but unfortunately do not fully eliminate them. For example, Amazon and third parties can still use the

metadata generated through the voice commands to profile users, and then serve them targeted advertisements. Further, the metadata could also be shared with other parties, a common practice in several IoT platforms [36, 40, 41, 53, 54, 64, 66]. These practices are particularly concerning in context of third parties because neither users nor Amazon have any visibility or control over the processing, sharing, and selling of user data; worse, third party skills often do not publish their privacy policies, nor adhere to them even when they do [30].

Thus, in this paper, we focus on the collection, usage, and sharing of metadata, that is generated through installing and interacting with skills, and refer to it as *smart speaker interaction data* or *user data*, throughout the paper.

2.3 Research questions

To the best of our knowledge, prior work lacks an in-depth analysis of the collection, usage, and sharing of user data in smart speaker ecosystems. To fill this gap, we systematically analyze the data collection, usage, and sharing practices in Alexa Echo smart speaker platform, including third party skills. We conduct controlled experiments where we intentionally expose data by installing and interacting with skills and observe platform’s behavior from three perspectives: (i) *network traffic* exchanged by smart speakers, (ii) *advertisements* served to smart speaker users, and (iii) *privacy policies* published by third party skills. Our goal is to combine these perspectives to answer the following research questions.

RQ1: Which organizations collect and propagate user data? We use the remote endpoints of network traffic to measure data collection and sharing by Amazon and third party skills. We can intercept and observe communication between Amazon and some (but not all) third parties; however, the Amazon ecosystem does not provide interfaces to facilitate comprehensive analysis of data collection, usage, and sharing. This motivates the need for inference below.

RQ2: Is smart speaker interaction data used by either Amazon or third party skills beyond purely functional purposes, such as for targeted advertising? We measure advertisements to infer data usage and sharing by Amazon and third party skills. To this end, we focus on detecting behaviorally targeted web ads. We study targeting in web ads because web publishers almost universally employ well-established programmatic advertising protocols [37, 38].¹

RQ3: Are data usage and sharing practices compliant with privacy policies and other disclosures? We extract key elements from privacy policies of Amazon and third party skills. For Amazon, we also review Alexa specific policies in Alexa Privacy Hub [6] and Alexa Device FAQs [2]. We also compare policy disclosures with our measurements and inferences to assess the compliance of data collection, usage, and sharing.

3 AUDITING FRAMEWORK

In this section, we describe our methodology for measuring data collection, usage, and sharing of user interaction data by Amazon and third party skills. Figure 1 presents the overview of our approach. At a high level, (1) we intentionally expose data by installing and

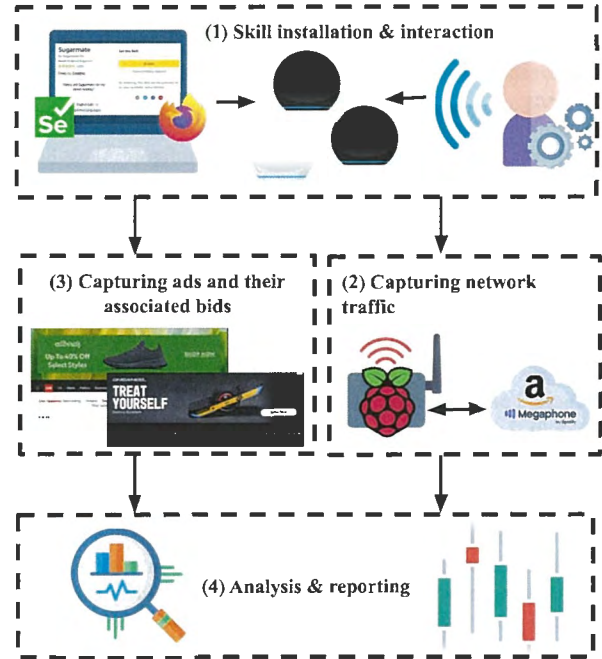


Figure 1: Approach overview: (1) We install and interact with skills from 9 different categories on 9 different smart speakers to train 9 smart speaker interest personas. (2) While installing and interacting, we also capture and store network traffic to/from Alexa Echo. In addition to interest (treatment) personas, we also train a vanilla (control) persona where we do not install and interact with skills. (3) We then visit popular websites while logged into each persona’s Amazon account to capture and store ads and their associated bids targeted to the personas. (4) We then analyze this data to measure data collection, usage, sharing, and its compliance with Amazon’s and skills policies.

interacting with skills on Alexa Echos; (2) capture the endpoints that collect this data by intercepting network traffic; and (3) then analyze ad targeting on popular websites to infer the usage of exposed data. These steps are described next, in Sections 3.1, 3.2, and 3.3, respectively. Our experiments were conducted from University of Washington and University of California, Davis in the US. All of the IP addresses for respective experiments geolocated to the same approximate location.

3.1 Simulating personas to expose data

3.1.1 Simulating interest personas. We simulate nine interest (treatment) personas by installing and interacting with skills from nine different Alexa marketplace categories: Connected Car, Dating, Fashion & Style, Pets & Animals, Religion & Spirituality, Smart Home, Wine & Beverages, Health & Fitness, and Navigation & Trip Planners. We simulate several personas because the nature

¹We do not directly study ads served on Echos because the Alexa advertising ecosystem is relatively nascent with several restrictions. For example, Amazon only allows audio ads on streaming skills [4] and typically requires rather high minimum ad spend commitment from advertisers [13].

IMC '23, October 24–26, 2023, Montreal, QC, Canada

Umar Iqbal et al.

of tracking, profiling, and ad targeting might differ across different skill categories. We refer to each interest persona according to its corresponding skill category name, and collectively as *interest personas*.

Skill installation. As a first step, we create dedicated Amazon accounts for each persona and use them to configure Alexa Echos (4th generation Alexa Echo smart speakers). To avoid contamination across personas, we configure each device through a fresh browser profile and assign it a unique IP address (all IPs geolocate to the same location). We then use a Selenium-based [69] web crawler to programmatically visit the Alexa skill marketplace, and iteratively install and enable the top-50 skills (based on the number of reviews) for each category. We use the dataset released in [30] to extract top skills from each category. If prompted, we enable all of the requested permissions by a skill. Note that we do not link accounts for skills that require it, since this is infeasible at scale in our testbed. Linking typically requires creating an account for the online service corresponding to the tested skill, and linking it to a physical IoT device, e.g., *iRobot* skill requires to link a robot vacuum cleaner with the skill [46].

Skill interaction. After installing each skill, we interact with it by programmatically uttering sample invocations listed by each skill.² We also parse skill descriptions to extract additional invocation utterances provided by the skill developer. We interact with the Alexa Echo by iteratively uttering each skill's invocations. In case Alexa expects a follow up response³ or has a response of more than 30 seconds, e.g., playing a song, we terminate the interaction by uttering *Alexa, Stop*.

3.1.2 Simulating control persona. In addition to the nine interest (treatment) personas, we also simulate a control persona, referred to as *vanilla* persona. Similar to interest personas, the vanilla persona is linked to an Amazon account, an Alexa Echo, and a unique IP address. However, we do not install or interact with skills on the vanilla persona. Vanilla persona serves as a baseline and allows to associate the deviation in the interest personas to the treatment, i.e., installation and interaction with skills, applied to them.

3.2 Capturing network traffic to measure data collection

We capture outgoing and incoming network traffic, to and from, Alexa Echos to measure data collection by Amazon and skills during skill installation and interaction. Alexa Echos do not support on-device network analysis, so we intercept network traffic on the router to capture the endpoints that collect data. To this end, we set up a custom Raspberry Pi (RPi) based router [63] to intercept incoming and outgoing network traffic. We uninstall each skill before installing the next one, to ensure that we associate the correct network traffic to each skill. Note that the captured network traffic is end-to-end encrypted, so we can inspect the *endpoints* that collect data, but not the *plaintext payloads*.

²We use *gTTS*, a python package for Google Translate text-to-speech API [28] to convert the textual commands to utterances.

³If Alexa expects a follow up response and the response is not provided, Alexa asks for the response a few times. We match the last two responses to determine if a follow up response is expected.

Inferring destination. The captured network traffic contains the IP addresses of contacted endpoints. We resolve these IP addresses to domain names by using the information from Domain Name System (DNS) packets in network traffic. We further map domain names to their parent organization by leveraging information from DuckDuckGo [27], Crunchbase [25], and WHOIS.

3.3 Capturing advertisements to infer data usage and sharing

We rely on ad content and advertisers' bidding behavior to infer data usage and sharing. Ad content can reveal the ad topic and consequently the user interests that advertisers might have inferred from the exposed Alexa Echo interaction data. However, ad content may lack objective or discernible association with the exposed data. For example, active advertising campaigns at the time of our experimentation may lack apparent association with the exposed data or advertising models may interpret user interests differently. In part to offset subjectivity, we use advertisers' bidding behavior to infer the usage and sharing of smart speaker interaction data. Prior research [24, 59, 60] has shown that the advertisers' bidding behavior is influenced by their pre-existing knowledge of the users, which typically results in higher bid values than cases where such knowledge is absent. Thus, if we encounter high bid values from advertisers, a likely cause is the usage and sharing of Alexa Echo interaction data.

Web advertisements. The header bidding protocol [37] exposes bid values to client-side browser, so we collect ad bids and ad images on sites where header bidding is supported, both after skill installation and skill interaction. To this end, we first identify top websites that support *prebid.js* [62], the most widely used implementation of header bidding protocol [48], and then visit those websites to capture bids and ad images. We extend OpenWPM [31] to identify and capture data on *prebid.js* supported websites. To identify *prebid.js* supported websites, we crawl Tranco top websites list (from 09/07/2021) [49] and probe for *prebid.js* version. We treat a website as *prebid* supported if we receive a non-null *prebid.js* version. We stop the crawl as soon as we identify 200 *prebid* supported websites. We then crawl the *prebid.js* supported websites and intercept bidding requests (or request bids if no bid requests are made).

To more accurately simulate user behavior, we enable OpenWPM's bot mitigation and wait for 10–30 seconds between webpages. We also visit the webpages in browser's native mode, with window size of 1366 × 678, instead of using the headless mode. Note that we crawl the *prebid.js* supported websites using the same browser profiles that are logged into the Amazon account and Alexa web companion app, and same IP addresses used to configure interest and vanilla personas (Section 3.1). The browser profiles and IP addresses connect personas with browsers and allow us to collect the advertisements targeted to the personas. To make sure that we visit websites with unchanged browser profiles and to not let webpages influence profiles, we ignore the updates made by the webpages to the browser profiles.

Interpreting bids. In addition to user interests, advertisers consider several factors, e.g., day of the week, website popularity, to determine the bid values [59, 60]. To reduce variability due to such

confounding factors, we strive to keep conditions consistent across personas. Specifically, we use identical hardware/software, collect bids at the same time (simultaneously), from the same location, and on the same websites, for all personas. In addition, we consider only bids from ad slots that are successfully loaded across all personas, because bid values vary by ad slots [60] and advertisers may not bid on all ad slots across all personas. We compare relative bid values across control and interest (treatment) personas because the absolute values can change over time, e.g., travel advertisements may get higher bids around holidays. Since it is non-trivial to reverse engineer and control for all the factors incorporated by advertisers, we crawl and extract bids from the `prebid.js` supported websites several times, i.e., 6 times before interacting with skills and 25 times after interacting with skills⁴, to further account for the variability in bid values.

Capturing requests/responses. In addition to collecting ad bids and images, we also record the network requests and responses while crawling popular websites. Network traffic allows us to measure data sharing (e.g., cookie syncing [39]) between Amazon and its advertising partners. Note that the network traffic captured while crawling (referred to as *web traffic*) is separate from the network traffic captured from Alexa Echos (Section 3.2).

4 NETWORK TRAFFIC ANALYSIS

In this section, we analyze network traffic from the Alexa Echos. We identify online services that directly collect user data and also investigate the functionality offered by these services.

4.1 Amazon as skill mediator

We first analyze network traffic to identify the services that collect user data. Table 1 presents the list of domains contacted by skills. We note that unlike more established platforms, e.g., mobile, the majority of the traffic in Alexa smart speaker ecosystem goes to the device manufacturer. Specifically, 446 (99.11%), 2 (0.45%), and 31 (6.89%) of the skills contact domains that belong to Amazon, skill vendors, and third parties, respectively. Four (0.89%) skills failed to load. All active skills contact Amazon because Amazon mediates communication between skills and users, i.e., Amazon first interprets the voice input and then shares it with the skill [17]. Another possible explanation for this is the use of Amazon’s infrastructure to host skills [5]. *Garmin* [33] and *YouVersion Bible* [81] are the only skills that send traffic to their own domains.

Figure 2 shows the number of network flows from skills to domains, their functionality, and their parent organizations. Similar to results from Table 1, most network flows involve Amazon. We note that the skills in most categories, except for Smart Home, Wine & Beverages, Navigation & Trip Planners, contact third party services.

4.2 Data shared with advertisers & trackers

We next analyze the functionality offered by services that collect user data. Several domains contacted by skills offer audio advertising and tracking services (rows highlighted in gray in Table 1).

⁴We terminated the experiment after 6 iterations after skill installation because we did not notice any personalization. We continued to crawl 25 times after skill interaction because we noticed personalization (Section 5) and needed more samples due to the variability in bid values.

Org.	Domains	Skills
Amazon	*(11).amazon.com	895
	prod.amcs-tachyon.com	305
	api.amazonalexa.com	173
	*(7).cloudfront.net	146
	device-metrics-us-2.amazon.com	123
	*(4).amazonaws.com	52
	acsechocaptiveportal.com	27
	fireoscaptiveportal.com	20
	ingestion.us-east-1.prod.arteries.alexa.a2z.com	7
	ffs-provisioner-config.amazon-dss.com	2
skill vendor	*(2).youversionapi.com	2
	static.garmincdn.com	1
Third party	dillilabs.com	9
	*(2).megaphone.fm	9
	cdn2.voiceapps.com	7
	*(2).podtrac.com	7
	*(2).pod.npr.org	4
	chtbl.com	3
	1432239411.rsc.cdn77.org	3
	*(2).libsyn.com	3
	*(3).streamtheworld.com	3
	discovery.meethue.com	2
	turnernetworksales.mc.tritondigital.com	1
	traffic.omny.fm	1

Table 1: Amazon, skill vendors, and third party domains contacted by skills. “Org.” column refers to organization. “Skills” column represents the count of skills. Advertising and tracking domains are shaded with grey. Subdomain counts are represented with *(#), e.g., *(11).amazon.com represents requests to 11 subdomains of amazon.com.

Organization	Functional	Advertising & Tracking	Total
Amazon	90.04%	6.80%	96.84%
Skill vendor	0.17%	0%	0.17%
Third party	1.49%	1.50%	2.99%
Total	91.70%	8.3%	100%

Table 2: Distribution of advertising / tracking and functional network traffic by organization.

We rely on filter lists [77] and manual analysis to detect advertising and tracking services. Table 2 provides the distribution of functional and advertising domains contacted by skills. We note that 9.4% of all network traffic, including 1.5% third party network traffic, supports advertising and tracking functionality. We note that device-metrics-us-2.amazon.com, used by Amazon to collect device metrics [20], is the most prominent tracking domain. The most contacted third party advertising and tracking services include Megaphone (megaphone.fm) and Podtrac (podtrac.com), both of which specialize in audio advertising and tracking services. We note that prominent skills, such as *Genesis* [35] and *Men’s Finest Daily Fashion Tip* [55] with 398 and 13 reviews, contact third party

IMC '23, October 24–26, 2023, Montreal, QC, Canada

Umar Iqbal et al.

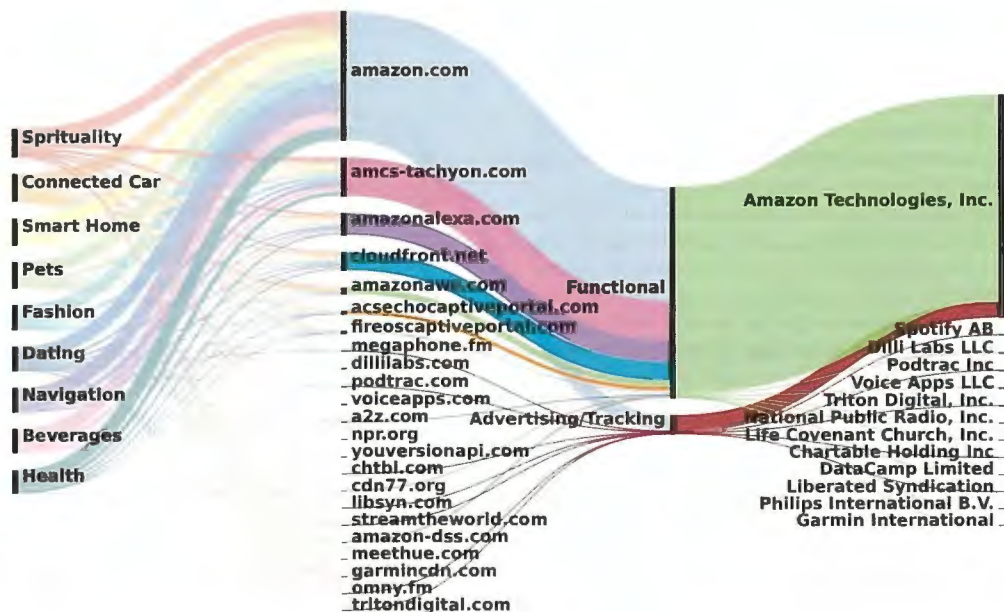


Figure 2: Network traffic distribution by persona, domain name, purpose, and organization.

Persona	Advertising & Tracking	Functional
Fashion & Style	9	4
Connected Car	7	0
Pets & Animals	3	11
Religion & Spirituality	3	10
Dating	5	1
Health & Fitness	0	1

Table 3: Count of advertising/tracking and functional third party domains contacted by personas.

advertising and tracking services. Despite Amazon’s Alexa advertising policy restricting non-streaming skills from playing ads [4, 18], we find that six non-streaming skills contact third party advertising and tracking services. Surprisingly, we note that these skills do not play any advertisements, despite including advertising services. It is unclear as to why non-streaming skills include advertising and tracking services and whether such skills should be flagged during skill certification [14].

Tables 3 and 4 further provide the distribution of advertising and tracking domains by personas and skills. From Table 3, we note that skills in five personas contact third party advertising and tracking services, with skills in the Fashion & Style persona contacting the most advertising and tracking services. Table 4 shows that skills contact several advertising and tracking services. For example, the *Garmin* [33] skill contacts as many as four advertising and tracking services.

Skill name	Advertising & Tracking
Garmin [33]	chtbl.com
	traffic.omny.fm
	dts.podtrac.com
	turnernetworksales.mc.tritondigital.com
Makeup of the Day [79]	*(2).megaphone.fm
	play.podtrac.com
	chtbl.com
Men's Finest Daily Fashion Tip [55]	play.podtrac.com
	*(2).megaphone.fm
Dating and Relationship Tips and advices [1]	play.podtrac.com
	*(2).megaphone.fm
Charles Stanley Radio [43]	*(2).streamtheworld.com

Table 4: Top-5 skills that contact third party advertising and tracking services. Subdomain counts are represented with *(#), e.g., *(2).megaphone.fm represents two subdomains of megaphone.fm.

5 AD TARGETING ANALYSIS

In this section, we analyze whether collected data is used to profile users, as well as infer if that profiling is used in ad targeting.

5.1 Interactions are used to infer interests

Since Amazon allows users to access data collected about them, we request data for interest and vanilla personas [11]. The data contains detailed information about device diagnostics, search history, retail interactions, Alexa, advertising, and other Amazon services.

Configuration	Persona	Amazon inferred interests
Installation	Health & Fitness	Electronics Home & Garden: DIY & Tools
		Home & Garden: DIY & Tools
Interaction (1)	Health & Fitness	Home & Garden: DIY & Tools
	Fashion & Style	Beauty & Personal Care Fashion Video Entertainment
		Electronics Home & Garden: DIY & Tools Home & Garden: Home & Kitchen
	Smart Home	Electronics Home & Garden: DIY & Tools Home & Garden: Home & Kitchen
Interaction (2)	Fashion & Style	Fashion Video Entertainment
	Smart Home	Pet Supplies Home & Garden: DIY & Tools Home & Garden: Home & Kitchen

Table 5: Advertising interests inferred by Amazon for interest personas. Installation represents advertising interest inferred after skill installation. Interaction represent advertising interests inferred after skill interaction. Data is downloaded twice after interaction, represented by (1) and (2).

We are mostly interested in advertising interests inferred by Amazon based on skill installation and interactions. We request data thrice, once after skill installation and twice after skill interaction, to see the evolution in inferred interests over time. Since advertising interests are inferred instantly and made available to users for download within 2 days [83], we request user interests after 3 days of skill installation and 8 and 31 days of skill interaction. Amazon on average took around 12 days to return the inferred interests after our request.

Table 5 presents the advertising interests inferred by Amazon for Health & Fitness, Fashion & Style, and Smart Home personas. For remaining personas, Amazon did not return any interests. We note that both skill installation and interaction lead to interests inference by Amazon. With only skill installation, Amazon infers that Health & Fitness persona is interested in *Electronics* and *DIY & Tools*. Skill interaction further allows Amazon to infer interests for Fashion & Style and Smart Home personas and also refine interests for Health & Fitness persona. Some of the interests inferred by Amazon seem clearly relevant to the personas. For example, *Fashion* and *Beauty & Personal Care* interests seem relevant to the Fashion & Style persona and *Home & Kitchen* interests seem relevant to the Smart Home persona. Note that for our second data request after interaction, Amazon did not return advertising interest files for Health & Fitness, Wine & Beverages, Religion & Spirituality, Dating, and vanilla personas—nor did they provide these files upon a third request.

Our results suggest that Amazon at the very least uses the meta-data of interactions with Alexa Echo smart speakers to infer user interests for ad targeting. This is concerning because before our research and public disclosure, these practices were not clearly stated in Amazon’s policies [3, 7] (see Section 6.2 for details).

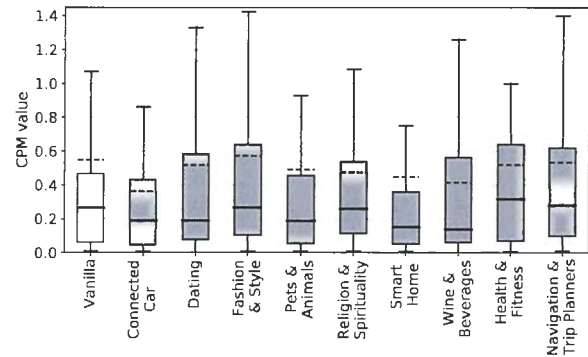


Figure 3: CPM values across vanilla (control) and interest (treatment) personas on common ad slots with skill installation. Solid and dotted lines in bars represent median and mean, respectively. Vanilla persona does not involve skill installation.

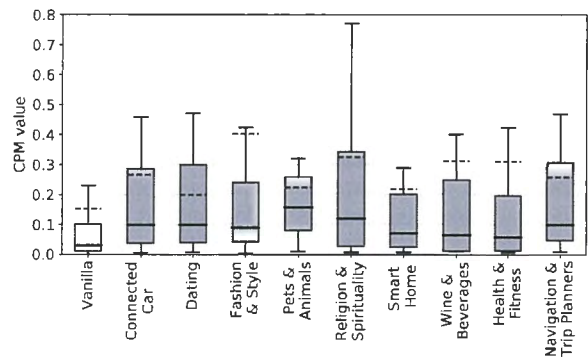


Figure 4: CPM values across vanilla (control) and interest (treatment) personas on common ad slots with user interaction. Solid and dotted lines in bars represent median and mean, respectively. Vanilla persona does not involve user interaction.

5.2 Interactions lead to higher ad bids

5.2.1 Bid values after skill installation. To evaluate whether only skill installation leads to personalized ad targeting, we first analyze advertisers bidding behavior for vanilla and interest personas without any user interaction. Figure 3 presents bid (CPM)⁵ values (y-axis) across vanilla and interest personas (x-axis) on common ad slots without user interaction. The figure shows that without user interaction, there is no discernible difference between vanilla and interest personas. Only Fashion & Style personas has slightly higher, statistically insignificant, mean bid value than vanilla persona. Overall, skill installation (without user interaction) does not lead to observable targeting.

⁵CPM (cost per mille) is the amount an advertiser pays a website per thousand visitors who see its advertisements. Bids are expressed in CPM.

Persona	Median	Mean
Connected Car	0.099	0.267
Dating	0.099	0.198
Fashion & Style	0.090	0.403
Pets & Animals	0.156	0.223
Religion & Spirituality	0.120	0.323
Smart Home	0.071	0.218
Wine & Beverages	0.065	0.313
Health & Fitness	0.057	0.310
Navigation & Trip Planners	0.099	0.255
Vanilla	0.030	0.153

Table 6: Median and mean bid values (CPM) for interest (treatment) and vanilla (control) personas with user interaction. Vanilla persona does not involve user interaction.

5.2.2 Bid values after skill installation and interaction. Next, we analyze advertisers bidding behavior for vanilla and interest personas with user interaction to evaluate if interaction with skills leads to personalized ad targeting. Figure 4 presents bid (CPM) values across vanilla and interest personas on common ad slots with user interaction. In contrast to bid values without user interaction (Figure 3), with user interaction (Figure 4) the bid values are significantly higher for interest personas as compared to vanilla persona. Table 6 shows the median and mean bid values for interest and vanilla personas with user interaction. The table indicates that median bids for all interest personas, except for Health & Fitness, are 2× higher than vanilla persona. Similarly, mean bids for four interest personas, i.e., Fashion & Style, Religion & Spirituality, Wine & Beverages, and Health & Fitness, are 2× higher than vanilla persona. We note that the bid values for Health & Fitness and Fashion & Style go as much as 30× and 27× higher than the mean of vanilla persona. It can also be seen from Table 6 that the mean bid values are higher than the median bid values, suggesting that some advertisers bid much higher than others. One possible explanation for this behavior could be that some advertisers have more information about the persona’s interests than the others, which leads them to place much higher bids than others. We discuss high absolute bid values with just skill installation in Appendix A.

5.2.3 After user interaction, interest personas receive significantly higher bids. We perform Mann-Whitney U test to analyze whether interest personas after user interaction receive significantly higher bids than vanilla persona. Since we perform multiple comparisons, we adjust our statistical significance tests with the Holm-Bonferroni correction method. Our null hypothesis is that the bid distributions for interest personas are similar to vanilla persona. Our alternative hypothesis is that the bid distributions for interest personas are higher than the vanilla persona. We reject the null hypothesis when the p -value is less than 0.05. In addition to p -value, we also report the effect size (rank-biserial coefficient). Effect size ranges from -1 to 1, where -1, 0, and 1 indicate stochastic subservience, equality, and dominance of interest persona over vanilla persona. Effect size between 0.11–0.28, 0.28–0.43, and ≥ 0.43 are considered small, medium, and large, respectively.

Persona	p -value	Adjusted p -value	Effect size
Connected Car	0.003	0.026	0.354
Dating	0.006	0.030	0.363
Fashion & Style	0.010	0.040	0.319
Pets & Animals	0.005	0.030	0.428
Religion & Spirituality	0.004	0.030	0.356
Smart Home	0.075	0.225	0.210
Wine & Beverages	0.083	0.225	0.192
Health & Fitness	0.149	0.225	0.139
Navigation & Trip Planners	0.002	0.015	0.410

Table 7: Statistical significance between vanilla (control) and interest (treatment) personas. p -value is computed through Mann-Whitney U test and adjusted through Holm-Bonferroni method. Effect size is rank-biserial coefficient. Personas with p -value less than 0.05 are shaded with grey.

Table 7 presents the results of statistical significance tests. We note that six interest personas have significantly higher bids than vanilla persona with medium effect size. For the remaining three interest personas, i.e., Smart Home, Wine & Beverages, and Health & Fitness, the differences are not statistically significant.

5.2.4 After user interaction, interest personas are targeted with personalized ads. Next, we analyze the ads delivered through prebid.js to personas after user interaction. In total, we receive 20,210 ads across 25 iterations. Since ads may lack any objective or even discernible association with the shared interests, as discussed in Section 3.3, we resort to manual analysis of ads. However, manual ad analysis is a tedious task and it is not feasible to analyze thousands of ads. To this end, we only manually analyze ads from Amazon and ads from installed skill vendors in their respective personas (e.g., an ad from Ford in Connected Car persona because it contains the *FordPass* skill) because we expect these ads to be the most personalized. We consider an ad to be personalized if it is only present in one persona and references a product in the same industry as the installed skills (e.g., an ad for a vehicle is shown to the Connected Car persona). While any manual labeling process is subject to human error and subjectivity, we argue that our definition is sufficiently concrete to mitigate these concerns.

In total, we filter 79 ads from installed skills’ vendors in their respective personas and 255 ads from Amazon ads for manual analysis. Out of the 79 ads from installed skills vendors, 60, 12, 1, and 1 are from Microsoft, SimpliSafe, Samsung, and LG in Smart Home persona, respectively. Out of the remaining 5, 3 are from Ford and 2 are from Jeep in Connected Car persona. However, none of the ads from installed skills vendors are exclusive to the personas where their skills are installed, which indicates that these ads do not reveal obvious personalization.

Ads from Amazon do seem to be personalized to personas. Table 8 presents the unique ads from Amazon that show apparent personalization. Health & Fitness and Smart Home personas receive unique ads with apparent personalization, whereas Religion & Spirituality and Pets & Animals receive unique ads but without any apparent personalization. The dehumidifier ad (Figure 5a) appears to have an association with the *Air Quality Report* skill [42]

Tracking, Profiling, and Ad Targeting in the Alexa Echo Smart Speaker Ecosystem

IMC '23, October 24–26, 2023, Montreal, QC, Canada

Persona	Advertised products
Health & Fitness	Dehumidifier, Essential oils
Smart Home	Vacuum cleaner, Vac. clean. accessories
Religion & Spirituality	Wifi router, Kindle, Swarovski
Pets & Animals	PC files copying/switching software

Table 8: Ads from Amazon on interest personas. Green represents unique ads with apparent relevance to the persona. Yellow represents unique ads that repeat across iterations but do not have any apparent relevance to the persona.

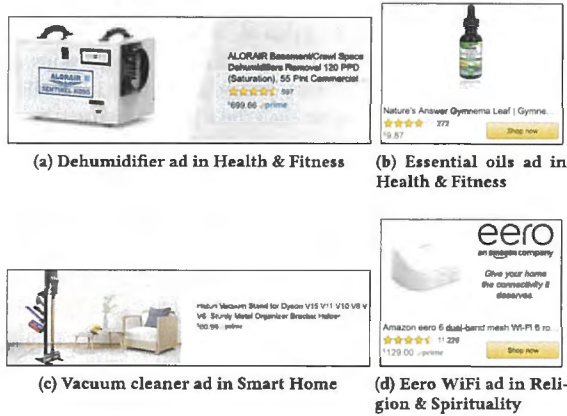


Figure 5: Unique and repeated ads in interest personas.

and the essential oils ad (Figure 5b) appears to have an association with the *Essential Oil Benefits* skill [74] in Health & Fitness persona. The dehumidifier ad appeared 7 times across 5 iterations and the essential oils ad appeared once in Health & Fitness persona. The vacuum cleaner and vacuum cleaner accessories ads (Figure 5c) from Dyson appear to have an association with the *Dyson* skill [29]; both ads appeared once in Smart Home persona. We notice several ads repeated across iterations in Religion & Spirituality and Pets & Animals personas that do not seem to have any apparent personalization. For example, Amazon Eero WiFi (Figure 5d), Amazon Kindle, and Swarovski ads exclusively appeared on 12, 14, 2 times across 8, 4, and 2 iterations, respectively in Religion & Spirituality persona. Similarly, PC files copying/switching software ad appeared 4 times in 2 iterations in Pets & Animals persona.

We also manually analyze ads from installed skill vendors and Amazon with only skill installation. We receive a total of 35 and 117 ads from skill vendors and Amazon with only skill installation, respectively. Out of the 35 ads from installed skills vendors, 16, 9, 2, and 2 are from Microsoft, SimpliSafe, Ring, and SharkNinja in Smart Home persona, respectively. Samsung, LG, and ATT also serve one ad each in Smart Home persona. Honda and Dodge serve one ad each in Connected Car and one Starbucks ad appeared in Wine & Beverages. However, none of the ads from installed skills vendors are exclusive to the personas where their skills are installed, which

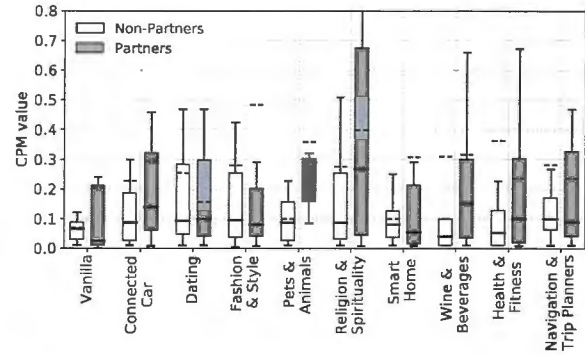


Figure 6: Bid values across personas on common ad slots distributed by Amazon's advertising partners.

indicates that these ads do not reveal obvious personalization. In case of Amazon, out of 117 ads, only two ads are unique to Health & Fitness persona, i.e., an ad for an electric toothbrush appearing once and an ad for an air fryer toaster appearing 4 times. However, similar to the ads from skill vendors, ads from Amazon also lack an apparent relevance to the personas as per our rubric, i.e., Health & Fitness persona does not have any skills related to electric toothbrush or an air fryer toaster oven.

Since we do not find a strong targeting signal in personas with only skill installation, we do not further analyze this case.

5.3 Sharing beyond the observed endpoints

Next, we infer the potential sharing of smart speaker interaction metadata from Amazon and third party skills, with other online services, not necessarily observable from the smart speaker.

5.3.1 Some advertisers sync their cookies with Amazon and bid higher than non-cookie syncing advertisers. To target personalized ads, advertisers share user data with each other. Typically, unique user identifiers, e.g., cookies, are shared at the client side with cookie syncing and user interest data is synced at the server side [21]. We analyze cookie syncing instances that involve Amazon advertising services in the web traffic captured while collecting ads (Section 3.3). We note that 41 third parties sync their cookies with Amazon across all Echo interest personas. Amazon did not sync its cookies with any advertiser.⁶ The one sided cookie-syncs could be explained by Amazon advertising's recent services for central identity resolution [78].

To infer potential data sharing by Amazon, we compare and contrast the bid values by Amazon's partners (i.e., cookie syncing advertisers) and non-partner advertisers. Figure 6 presents the bid values on common ad slots by Amazon's partner and non-partner advertisers. We note that the bids by partner advertisers are higher than that by non-partner advertisers on most personas. Table 9 shows the median and mean bid values by partner and non-partner advertisers. It can be seen from the table that both median and mean bid values from partners are high for 6 and 7 personas, respectively,

⁶We analyze the OpenWPM datasets released by prior work [45] to validate that Amazon's cookie syncing behavior is not unique to our dataset.

Persona	Partner		Non-partner	
	Median	Mean	Median	Mean
Connected Car	0.140	0.296	0.086	0.228
Dating	0.099	0.159	0.094	0.254
Fashion & Style	0.080	0.485	0.095	0.281
Pets & Animals	0.290	0.358	0.087	0.101
Religion & Spirituality	0.268	0.400	0.088	0.276
Smart Home	0.054	0.307	0.080	0.101
Wine & Beverages	0.150	0.316	0.041	0.310
Health & Fitness	0.099	0.235	0.053	0.363
Navigation & Trip Plan.	0.090	0.236	0.100	0.281
Vanilla	0.025	0.203	0.352	0.066

Table 9: Median and mean bid values for personas from Amazon’s partner and non-partner advertisers.

as compared to bids from non-partners. Median bid values are as much as 3× higher for Pets & Animals, Religion & Spirituality, and Wine & Beverages personas, while mean bid values are 3× higher for Pets & Animals, Smart Home, and vanilla personas. It is noteworthy that Amazon’s advertising partners further sync their cookies with 247 other third parties, including advertising services. Such cookie syncs may lead to the propagation of user data in the advertising ecosystem.

5.3.2 It is unclear whether skills play a role in targeting of personalized ads. We now discuss Amazon’s and skills’ role in higher bids and targeting of personalized ads. Since all interactions are mediated through Amazon, Amazon has the best vantage point to infer personas’ interests and target personalized ads. Specifically, all user commands are interpreted by Amazon and most network requests are routed to/through Amazon (Table 1 and Figure 2). Further, when a persona is logged into its Amazon account, Amazon can access its cookies during web visits. In fact, Sections 5.1 and 5.2.4 already show that Amazon infers users’ advertising interests from the meta-data of their interaction with Echos and uses the inferred interests to target personalized ads to users. We also note that Smart Home, Wine & Beverages, and Navigation & Trip Planners, personas do not contact any non-Amazon services but still receive high bid values, as compared to vanilla persona. Amazon further infers discernible interests for the Smart Home and Fashion & Style personas (Table 5). These results suggest that Amazon plays a crucial, if not a sole, role in higher bids and targeting of personalized ads.

In contrast, skills can only rely on persona’s email address (if given permission), IP address (if skills contact non-Amazon web services directly), and by collaborating with Amazon, to reach to personas. Though we allow skills to access email address, we do not log in to any online services (except for Amazon), thus skills cannot use email addresses to target personalized ads. Skills that contact non-Amazon web services and skills that collaborate with Amazon can still target ads to users. However, we note that only a handful of skills contact few advertising and tracking services (Table 1 and Figure 2). Similarly, we note that none of the skills re-target ads to personas (Section 5.2.4), which implies that Amazon might not be engaging in data sharing partnerships with skills.

6 COMPANIES’ REPRESENTATIONS

Our auditing framework, so far, measured, directly or indirectly, the actual practices of Amazon as well as skills, with respect to data collection, usage, and sharing. Companies also make representations and public disclosures, which should accurately and fully disclose their practices. Such disclosures include press releases, statements on the website, and the – legally binding – privacy policies. In Section 6.1, we focus on and analyze the consistency between the data collection practices of skill vendors (as directly observed in the network traffic) with the statements made in their privacy policies.

6.1 Privacy policy analysis

6.1.1 Collecting privacy policies. We download skills’ privacy policies from the *Developer Privacy Policy* link on the skill installation page. Recall from Section 3.1.1 that we experiment with 450 skills, i.e., top-50 skills from nine categories. Among the 450 skills, only 214 (47.6%) skills provide privacy policy links on their installation pages. The percentage is higher than the statistics reported by prior work [50], which identified that only 28.5% of the skills provide a privacy policy link. We surmise that it could be because we investigate popular skills. Unfortunately, only 188 skills out of 214 provide a valid privacy policy link. Further, among the 188 obtained privacy policies, 129 do not even mention the word “Alexa” or “Amazon” in their text. We manually read many of the privacy policies, and notice that they are mostly generic and apply to several products from the same developer. Thus, they do not seem to be specific to Alexa skills.

6.1.2 Network traffic flows are often inconsistent with privacy policies. As in prior work [19, 50, 73], we use on PoliCheck [19] to evaluate the consistency of network traffic flows with privacy policies. PoliCheck extracts $\langle \text{data type}, \text{entity} \rangle$ tuples from the network traffic and the textual disclosures in the privacy policies, and checks the consistency of the two. However, such analysis requires access to unencrypted network traffic, which is unavailable in our case (see Section 3.2). Thus, we adapt PoliCheck to perform the analysis only on the endpoints found in the encrypted traffic collected from the Alexa Echo.

Specifically, we modify PoliCheck to only validate the consistency of endpoint organizations contacted by skills with their privacy policies. We update PoliCheck’s entity ontology by inspecting the network traffic and including observed endpoints, which we then map to their organization using the methodology described in Section 3.2. Based on the service offered by the organization, it is assigned one or more categories from *platform provider*, *voice assistant service*, *analytic (tracking) provider*, *advertising network*, and *content provider*. These categories are derived from PoliCheck’s entity ontology and terms found in the privacy policies. We visit the website of each organization to determine the service offered by it. *Platform provider* and *voice assistant service* labels are only assigned to Amazon. We also update PoliCheck’s consistency disclosure definitions. Specifically, data flows are referred to as (1) *clear*, when the endpoints are disclosed in the privacy policy using the exact organization name; (2) *vague*, when the endpoint is disclosed using category names or *third party*; and (3) *omitted*, when the endpoint is not disclosed at all. We do not use *ambiguous* and *incorrect* disclosures because a contradiction cannot be determined without

considering data types. We label an endpoint as *no policy* when the skill does not provide a privacy policy.

Disclosure of platform-party collection. Table 10 presents the result of our endpoint analysis. The table shows that only 10 privacy policies clearly indicate that personal information is collected by Amazon. For example, the *Sonos* skill [70] clearly states that voice recordings are collected by Amazon. Furthermore, we find that 136 skills vaguely disclose that their network traffic may go to Amazon. For example, the *Harmony* skill [51] privacy policy mentions sending data but without referring to the name of the entity: “Circle products may send pseudonymous information to an analytics tool, including timestamps, transmission statistics, feature usage, performance metrics, errors, etc.”

Disclosure of first-party collection. We find that 32 skills connect to non platform-party endpoints. Among them, 10 provide privacy policies and only six have at least one clear or vague disclosure. The only two clearly disclosed first-party endpoints are in the privacy policies of the *YouVersion Bible* [81] and *Garmin* [33] skills: they correspond to the organizations that are the developers of the skills.

Disclosure of third party collection. Many skills rely on third party organizations, e.g., Liberated Syndication, Podtrac, Spotify and Triton Digital, which provide audio content distribution and tracking/advertising services. However, only a few skills disclose data collection and sharing with third party organizations in their privacy policies, and when they do, they use vague terms. For example, the *Charles Stanley Radio* skill [43] uses the term “external service providers” to refer to third party organizations in its privacy policy. Another example is the *VCA Animal Hospitals* skill that uses the blanket term “third parties” to refer to all third party organizations in its privacy policy [76].

6.2 Alexa Echo data processing policies

The Alexa Echo smart speaker data collection practices are clearly stated in Amazon’s privacy policy, i.e., Amazon collects data when users “talk to or otherwise interact with our Alexa Voice service” [12]. However, we did not find similar information about the usage of the Alexa Echo interaction data for user interest inference for ad targeting. We also explored Amazon’s dedicated Alexa specific policies, i.e., Alexa Privacy Hub [7] and Alexa Device FAQs [3], but similar to privacy policy, we did not find any information about the usage of Alexa Echo interaction data for ad targeting, at the time of our research. However, after our work’s preprint was released and Amazon was made aware [23], Amazon updated the Alexa Privacy Hub [6] and the Alexa Device FAQs [2] to include that Alexa Echo interaction data is used for ad targeting.

7 DISCUSSION

7.1 Parallels with other IoT platforms

7.1.1 Related platform-agnostic IoT works. Several IoT works have measured network traffic to detect data collection and sharing. For example, [41, 53, 54, 64, 66] have shown that tracking is common in several IoT platforms, regardless of the presence of specific apps/skills. In contrast to prior work, our study identifies that Alexa Echo smart speakers contact previously unreported endpoints from Amazon, skills vendors, and third parties. For example, with respect to the endpoints reported in a 2021 study [53], we have observed 4

new Amazon domains (acsechocaptiveportal.com, amazon-dss.com, a2z.com, amazonalexa.com), 2 skills-specific endpoints (see *skills* row in Table 1) and 12 new third party endpoints (see *third party* row in Table 1). A possible explanation could be the change in Alexa Echo ecosystem since it was last studied, e.g., api.amazonalexa.com may have replaced api.amazon.com, which was no longer contacted.

7.1.2 Related platform-specific IoT works. Compared to prior work on smart TVs [57, 75] and VR headsets [73], we found less observable data-tracking activity on smart speakers. However, ad targeting on the web, specifically from partner advertisers, indicates that data sharing may still be happening. A possible explanation could be the server-side interfaces from smart speaker platform that expose data for advertising purposes.

7.2 Possible mitigations

7.2.1 Improved transparency and control for users. Smart speaker users may want to know what data is being collected, how that data is being used, and by whom. Our work suggests the need for greater transparency for users about the answer to these questions, as well as better control. Such transparency and control might come through a redesign of the platform itself (e.g., improved privacy-related UX, system-level enforcement with information flow control) or through external audits (such as with our framework) and external controls (either technical—e.g., network traffic filtering—and/or policy-based). For example, Alexa Echos could be equipped with a debugging interface [58]. Making such an interface available for developers and auditors could provide direct observations of data sharing. To limit tracking, a user might use software to selectively block network traffic that is not essential for the skill to work (e.g., using an approach similar to [53]).

7.2.2 Limiting user interaction data. To limit the sharing of data, one can offload the wake-word detection and transcription functions of the Alexa platform with offline tools such as [61, 65], and just send to the Alexa platform the transcribed commands using their textual API with no loss of functionality. Data sharing to only one vendor could also be limited by allowing users an option to install voice assistants from their preferred vendor, similar to apps on mobile devices.

7.3 Generalizability to other platforms

The modularity of our framework makes it suitable to be generalized across other smart speaker platforms and even other IoT platforms, such as smart TVs and AR/VR. The core idea behind our framework of exposing data and inferring its usage is universally applicable. However, other platforms may require implementation changes to various modules of our framework. For example, our data exposure module that exposes data by uttering voice commands may be readily deployable to other smart speaker platforms but not to the AR/VR platforms. On the other hand, our data usage inference module that infers the usage of exposed data through online advertising on the web may be readily deployable across other smart speaker and IoT platforms. Similarly, our network traffic capturing and our privacy policy modules could also be readily deployed to other platforms. We also envision that other platforms

IMC '23, October 24–26, 2023, Montreal, QC, Canada

Umar Iqbal et al.

Endpoint organization	Categories in the ontology	Contacted skills
Amazon Technologies, Inc.	analytic provider, advertising network, content provider, platform provider, voice assistant service	AAA Road Service, Salah Time, My Dog, My Cat, Outfit Check!, Pet Buddy, Rain Storm by Healing FM, Single Decade Short Rosary, Islamic Prayer Times, Sonos, 136 skills, 42 skills, 258 skills
Chartable Holding Inc	analytic provider, advertising network	Garmin, Makeup of the Day, My Tesla (Unofficial)
DataCamp Limited	content provider	Relaxing Sounds: Spa Music, Chiron My Dog, Calm My Cat
Dilli Labs LLC	content provider	VCA Animal Hospitals, EcoSmart Live, Dog Squeaky Toy, Relax My Pet, Car Sounds, Tynosaur Sounds, Hush Puppy, Calm My Dog, Calm My Pet
Garmin International	content provider	Garmin
Liberated Syndication	analytic provider, advertising network	Calm My Pet, Al's Dog Training Tips
National Public Radio, Inc.	content provider	Makeup of the Day, Men's Finest Daily Fashion Tip
Philips International B.V.	content provider	Say a Prayer, Angry Girlfriend
Podtrac Inc	analytic provider, advertising network	Garmin, Gwynnie Bee, Genesis, Men's Finest Daily Fashion Tip, Love Trouble, Makeup of the Day, Dating & Relationship Tips
Spotify AB	analytic provider, advertising network	Gwynnie Bee, Genesis, Dating and Relationship Tips and advices, Makeup of the Day, Men's Finest Daily Fashion Tip, Love Trouble
Triton Digital, Inc.	analytic provider, advertising network	Garmin, Charles Stanley Radio
Voice Apps LLC	content provider	Prayer Time, Charles Stanley Radio, Morning Bible Inspiration, Holy Rosary, meal prayer, Halloween Sounds, Bible Trivia
Life Covenant Church, Inc.	content provider	YouVersion Bible, Lords Prayer

Table 10: Endpoint organizations observed in the network traffic from skills run on the Alexa Echo: only 32 skills exhibit non-Amazon endpoints. Skills highlighted in green use the exact organization name in the statement that discloses data collection and sharing by the endpoint. Skills highlighted in yellow use third party or other vague terms. Skills highlighted in red do not declare the contacted endpoint at all. Skills highlighted in gray do not provide a privacy policy.

might also be able to extend our framework by incorporating platform specific components. For example, smart TV platforms have a mature video advertising ecosystem [56], which could be leveraged to strengthen the inference of exposed data.

7.4 Limitations

Our simulated interactions with Alexa Echo skills are potentially different from how real users would interact with the skills. For example, we utter a specific set of predefined commands and do not complete the conversation flows, e.g., by responding to the follow up questions by the skills. However, even with partial conversation flows, we are able to establish that smart speaker interaction metadata is used to infer user interests, which is then used for ad targeting. We expect that with more complete conversation flows, the extent of tracking, profiling, and ad targeting would potentially increase even further.

In order to be able to attribute the usage of exposed data to entities in the smart speaker ecosystem (i.e., the smart speaker platforms, skill vendors, and third parties), we do not expose data to other online services. However, in a realistic setting, users might expose their data (e.g., emails) to other online services, making it challenging to assess the role of individual services in using user's

data. Future work could explore mechanisms to disentangle the role of individual services when data is exposed to several online services.

We currently visually inspect and analyze ad content, which can lead to subjective assessment and also hinder reproducibility. Future work could address that limitation by automating ad content analysis (e.g., with the help of machine learning) or by leveraging crowdsourcing techniques used by prior research for ad content analysis [21, 82] to foster reproducibility.

7.5 Ethics & Disclosure

We visit websites to collect ads and their associated bids, which results in ad impressions and could potentially cause advertisers to lose some revenue. However, we only visit the minimal number of websites necessary to establish statistical confidence in our inferences, to limit the economic impact of our experiments. Note that this concern is common to all web measurement studies and we follow commonly accepted practices to minimize the impact of our measurements.

We did not directly disclose our findings to Amazon because traditional vulnerability disclosures assume overlooked issues, e.g., a security bug because of an implementation flaw. In our study, the

issues we identified seem to be part of the design of the (eco)system and the purpose of our study is to bring public transparency. In fact, after our work's preprint was released and Amazon was made aware [23], Amazon updated its disclosure to include that it uses smart speaker interaction data for ad targeting [16]. We have also shared our findings in a public forum at the Federal Trade Commission (FTC) [44].

8 CONCLUSION

In this paper, we audited data collection, usage, and sharing practices in the Alexa Echo smart speaker ecosystem. Our results indicated that Alexa Echo interactions were tracked by both Amazon and third parties. We also found that Amazon used Alexa Echo interaction data to infer user interests and then used those inferences for ad targeting, which was not clearly stated in Amazon's policies before our research and public disclosure. In many instances, skills did not clearly disclose their data collection practices in their privacy policies, did not provide any privacy policy, or did not reference the platform's privacy policy. Given these findings, there is a clear need for increased transparency—by using auditing tools such as ours—on the practices of smart speaker platforms and third parties operating on them. Our auditing framework and results may be useful to several stakeholders, including Amazon and skill developers (for internal privacy audits), policymakers (for crafting and effectively enforcing regulation), and users (as an incentive to guard their privacy using available tools).

We make our code and datasets publicly available at <https://privsec-research.github.io/alexaechos>.

ACKNOWLEDGEMENTS

We would like to thank Caelan MacArthur who contributed to the preliminary investigation of this work in summer 2021 as part of the CRA-WP Distributed Research Experiences for Undergraduates (DREU) at the University of California, Davis. We would like to thank Jan Sanislo from the University of Washington IT team for their help with setting up a WiFi network for connecting Alexa Echo devices. This work is supported in part by the National Science Foundation under awards CNS-1956393, CNS-1955227, CNS-2103439, CNS-2103038, CNS-2138139, CNS-2114230, CNS-1909020, Computing Research Association for the CIFellows 2021 Project under Grant CNS-2127309, the Northeastern University Future Faculty Fellowship (2021), and the Consumer Reports Innovation Lab.

REFERENCES

- [1] Aaron M. Spelling. 2022. Dating and Relationship Tips and advices. <https://www.amazon.com/dp/B07YCKFCFC>.
- [2] Amazon.com, Inc. 2022. Alexa and Alexa Device FAQs. <https://www.amazon.com/gp/help/customer/display.html?nodeId=201602230>.
- [3] Amazon.com, Inc. 2022. Alexa and Alexa Device FAQs (archived September 2022). <https://web.archive.org/web/20220901073936/http://www.amazon.com/gp/help/customer/display.html?nodeId=201602230>.
- [4] Amazon.com, Inc. 2022. Alexa Blogs: Advertising and Alexa. <https://developer.amazon.com/blogs/alexa/post/54c3a0f8-5b29-4071-acd7-2b832b860c83/advertising-and-alexa>.
- [5] Amazon.com, Inc. 2022. Alexa-hosted Skills. <https://developer.amazon.com/en-US/docs/alexa/hosted-skills/build-a-skill-end-to-end-using-an-alexa-hosted-skill.html>.
- [6] Amazon.com, Inc. 2022. Alexa Privacy Hub. <https://www.amazon.com/Alexa-Privacy-Hub/b?ie=UTF8&node=19149155011>.
- [7] Amazon.com, Inc. 2022. Alexa Privacy Hub (archived October 2022). <https://web.archive.org/web/20221010091208/http://www.amazon.com/b?node=19149155011>.
- [8] Amazon.com, Inc. 2022. Alexa Skill Certification Requirements. <https://developer.amazon.com/en-US/docs/alexa/custom-skills/certification-requirements-for-custom-skills.html>.
- [9] Amazon.com, Inc. 2022. Alexa Skills Policy Testing. <https://developer.amazon.com/en-US/docs/alexa/custom-skills/policy-testing-for-an-alexa-skill.html>.
- [10] Amazon.com, Inc. 2022. Alexa Skills Privacy Requirements. <https://developer.amazon.com/en-US/docs/alexa/custom-skills/security-testing-for-an-alexa-skill.html#25-privacy-requirements>.
- [11] Amazon.com, Inc. 2022. Amazon: Request Your Data. <https://www.amazon.com/gp/privacycentral/dsar/preview.html>.
- [12] Amazon.com, Inc. 2022. Amazon.com Privacy Notice. <https://www.amazon.com/gp/help/customer/display.html?nodeId=GX7NJQ4ZB8A1HFRNJ>.
- [13] Amazon.com, Inc. 2022. Audio Ads – Create audio advertising campaigns. <https://advertising.amazon.com/en-ca/solutions/products/audio-ads>.
- [14] Amazon.com, Inc. 2022. AVS Testing and Certification Process. <https://developer.amazon.com/en-US/docs/alexa/alexa-voice-service/product-testing-overview.html>.
- [15] Amazon.com, Inc. 2022. Configure Permissions for Customer Information in Your Skill. <https://developer.amazon.com/en-US/docs/alexa/custom-skills/configure-permissions-for-customer-information-in-your-skill.html>.
- [16] Amazon.com, Inc. 2022. Managing advertising preferences on Alexa. <https://www.amazon.com/b?node=98592480011>.
- [17] Amazon.com, Inc. 2022. Module 2: Design an Engaging Voice User Interface. <https://developer.amazon.com/en-US/alexa/alexa-skills-kit/get-deeper/tutorial-1s-code-samples/build-an-engaging-alexa-skill/module-2>.
- [18] Amazon.com, Inc. 2022. Policy Testing. <https://developer.amazon.com/en-US/docs/alexa/custom-skills/policy-testing-for-an-alexa-skill.html#advertising>.
- [19] Benjamin Andow, Samin Yaseer Mahmud, Justin Whitaker, William Enck, Bradley Reaves, Kapil Singh, and Serge Egelman. 2020. Actions Speak Louder than Words: Entity-Sensitive Privacy Policy and Data Flow Analysis with PoliCheck. In *29th USENIX Security Symposium (USENIX Security 20)*. USENIX Association, Boston, 985–1002. <https://www.usenix.org/conference/useenixsecurity20/presentation/andow>.
- [20] Rubén Barceló-Armada, Ismael Castell-Uroz, and Pere Barlet-Ros. 2022. Amazon Alexa traffic traces. *Computer Networks* 205 (2022), 108782.
- [21] Muhammad Ahmad Bashir, Sajjad Arshad, William Robertson, and Christo Wilson. 2016. Tracing Information Flows Between Ad Exchanges Using Retargeted Ads. In *25th USENIX Security Symposium (USENIX Security 16)*. USENIX Association, Austin, 481–496.
- [22] Long Cheng, Christin Wilson, Song Liao, Jeffrey Young, Daniel Dong, and Hongxin Hu. 2020. Dangerous Skills Got Certified: Measuring the Trustworthiness of Skill Certification in Voice Personal Assistant Platforms. In *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security (CCS '20)*. Association for Computing Machinery, Virtual, 1699–1716.
- [23] Thomas Claburn. 2022. Study: How Amazon uses Echo smart speaker conversations to target ads. https://theregister.com/2022/04/27/amazon_audio_data/.
- [24] John Cook, Rishab Nithyanand, and Zubair Shafig. 2020. Inferring Tracker-Advertiser Relationships in the Online Advertising Ecosystem using Header Hiding. In *Proceedings on Privacy Enhancing Technologies*, Vol. 2020 (1). Sciendo, Virtual, 65–82.
- [25] Crunchbase Inc. 2022. Crunchbase. <https://www.crunchbase.com/>.
- [26] Daniel J. DuBois, Roman Kolcun, Anna Maria Mandalari, Muhammad Talha Paracha, David Choffnes, and Hamed Haddadi. 2020. When Speakers Are All Ears: Characterizing Misactivations of IoT Smart Speakers. In *Proceedings on Privacy Enhancing Technologies*, Vol. 2020 (4). Sciendo, Virtual, 255–276.
- [27] DuckDuckGo. 2022. Tracker Radar (list of entities). <https://github.com/duckduckgo/tracker-radar/tree/main/entities>.
- [28] Pierre N. Durette. 2022. gTTS (Google Text-to-Speech), a Python library and CLI tool to interface with Google Translate text-to-speech API. <https://pypi.org/project/gTTS/>.
- [29] Dyson Limited. 2022. Dyson. <https://www.amazon.com/dp/B06WV7N7SHC>.
- [30] Jide S. Edu, Xavier Ferrer-Aran, Jose Such, and Guillermo Suarez-Tangil. 2023. SkillVet: Automated Traceability Analysis of Amazon Alexa Skills. *IEEE Transactions on Dependable and Secure Computing* 20, 1 (2023), 161–175.
- [31] Steven Englehardt and Arvind Narayanan. 2016. Online Tracking: A 1-Million-Site Measurement and Analysis. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS '16)*. Association for Computing Machinery, Vienna, 1388–1401.
- [32] Geoffrey A. Fowler. 2019. Alexa has been eavesdropping on you this whole time. <https://www.washingtonpost.com/technology/2019/05/06/alexa-has-been-eavesdropping-you-this-whole-time/>.
- [33] Garmin International. 2022. Garmin. <https://www.amazon.com/dp/B075TRB4V5>.
- [34] Gary Horcher. 2018. Woman says her Amazon device recorded private conversation, sent it out to random contact. <https://www.kiro7.com/news/local/woman-says-her-amazon-device-recorded-private-conversation-sent-it-out-to-random-contact/755507974/>.
- [35] Genesis Motors USA. 2022. Genesis. <https://www.amazon.com/dp/B01JXP09PI>.

IMC '23, October 24–26, 2023, Montreal, QC, Canada

Umar Iqbal et al.

- [36] Aniket Girish, Tianrui Hu, Vijay Prakash, Daniel J. Dubois, Srđjan Matic, Danny Yuxing, Serge Egelman, Joel Reardon, Juan Tapiador, David Choffnes, and Narseo Vallina-Rodriguez. 2023. In the Room Where It Happens: Characterizing Local Communication and Threats in Smart Homes. In *Proceedings of the ACM Internet Measurement Conference (IMC '23)*. Association for Computing Machinery, Montréal.
- [37] Google, Inc. 2022. Header Bidding. <https://admanager.google.com/home/resources/feature-brief-open-bidding/>.
- [38] Google, Inc. 2022. Real-time Bidding. <https://developers.google.com/authorized-buyers/rth/start>.
- [39] Google, Inc. 2022. RTB - Cookie Matching. <https://developers.google.com/authorized-buyers/rth/cookie-guide>.
- [40] Tianrui Hu, Daniel J. Dubois, and David Choffnes. 2023. BehavIoT: Measuring Smart Home IoT Behavior Using Network-Inferred Behavior Models. In *Proceedings of the ACM Internet Measurement Conference (IMC '23)*. Association for Computing Machinery, Montréal.
- [41] Danny Yuxing Huang, Noah Aphorpe, Frank Li, Gunes Acar, and Nick Feamster. 2020. IoT Inspector: Crowdsourcing Labeled Network Traffic from Smart Home Devices at Scale. In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 4. Association for Computing Machinery, New York, NY, USA, Article 46, 21 pages. <https://doi.org/10.1145/3397333>
- [42] ICM. 2022. Air Quality Report. <https://www.amazon.com/dp/B01E0FCHMA>.
- [43] In Touch Ministries. 2022. Charles Stanley Radio. <https://www.amazon.com/dp/B07FF2QGKW>.
- [44] Umar Iqbal. 2022. Your Echos are Heard: Tracking, Profiling, and Ad Targeting in the Amazon Smart Speaker Ecosystem, FTC PrivacyCon 2022. <https://www.ftc.gov/news-events/events/2022/11/privacycon-2022>.
- [45] Umar Iqbal, Charlie Wolfe, Charles Nguyen, Steven Englehardt, and Zubair Shafiq. 2022. Khaleesi: Breaker of Advertising and Tracking Request Chains. In *31st USENIX Security Symposium*. USENIX Association, Boston, 2911–2928.
- [46] iRobot. 2022. iRobot Home. <https://www.amazon.com/dp/H06Y3PSHQ3>.
- [47] Huaifeng Jin and Shuo Wang. 2018. Voice-based determination of physical and emotional characteristics of users. US Patent 10096319B1.
- [48] Kevel. 2022. Header Bidding (HBLX) 2021 Tracker. <https://www.kevel.com/hbix/>.
- [49] Victor Le Pochat, Tom Van Goethem, Samaneh Tajalizadehkhoob, Maciej Korczyński, and Wouter Joosen. 2019. Tranco: A Research-Oriented Top Sites Ranking Hardened Against Manipulation. In *26th Annual Network and Distributed System Security Symposium*. Internet Society, San Diego.
- [50] Christopher Lentzsch, Sheel Jayesh Shah, Benjamin Andow, Martin Degeling, Anupam Das, and William Enck. 2021. Hey Alexa, is this skill safe?: Taking a closer look at the Alexa skill ecosystem. In *28th Annual Network and Distributed System Security Symposium*. The Internet Society, San Diego.
- [51] Logitech. 2022. Harmony. <https://www.amazon.com/dp/B01M4LDPX3>.
- [52] Sapna Maheshwari. 2018. Hey, Alexa, What Can You Hear? And What Will You Do With It? <https://www.nytimes.com/2018/03/31/business/media/amazon-google-privacy-digital-assistants.html>.
- [53] Anna Maria Mandalari, Daniel J. Dubois, Roman Kolcun, Muhammad Talha Paracha, Hamed Haddadi, and David Choffnes. 2021. Blocking Without Breaking: Identification and Mitigation of Non-Essential IoT Traffic. In *Proceedings on Privacy Enhancing Technologies*, Vol. 2021 (4). Sciencio, Virtual, 369–388.
- [54] M. Hammad Mazhar and Zubair Shafiq. 2020. Characterizing Smart Home IoT Traffic in the Wild. In *2020 IEEE/ACM Fifth International Conference on Internet-of-Things Design and Implementation (IoTDI)*. IEEE, Sydney, 203–215.
- [55] Men's Finest. 2022. Men's Finest Daily Fashion Tip. <https://www.amazon.com/dp/B07CB3ZN6N>.
- [56] Hooman Mohajeri Moghaddam. 2022. Tracking and Behavioral Targeting on Connected TV Platforms. (2022). <https://dataspace.princeton.edu/handle/88435/dsp010p096b14c> Doctoral Dissertation.
- [57] Hooman Mohajeri Moghaddam, Gunes Acar, Ben Burgess, Arunesh Mathur, Danny Yuxing Huang, Nick Feamster, Edward W. Felten, Prateek Mittal, and Arvind Narayanan. 2019. Watching You Watch: The Tracking Ecosystem of Over-the-Top TV Streaming Devices. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security (CCS '19)*. Association for Computing Machinery, London, 131–147.
- [58] Tom Nardi. 2019. Uncovering the Echo Dot's Hidden USB Port. <https://hackaday.com/2019/08/15/uncovering-the-echo-dots-hidden-usb-port/>.
- [59] Lukasz Olejnik, Minh-Dung Tran, and Claude Castelluccia. 2014. Selling Off Privacy at Auction. In *21st Annual Network and Distributed System Security Symposium*. The Internet Society, San Diego.
- [60] Panagiotis Papadopoulos, Nicolas Kourtellis, Pablo Rodríguez Rodríguez, and Nikolaos Laoutaris. 2017. If You Are Not Paying for It, You Are the Product: How Much Do Advertisers Pay to Reach You?. In *Proceedings of the 2017 Internet Measurement Conference*. Association for Computing Machinery, London, 142–156.
- [61] Picovoice Inc. 2022. Porcupine Wake Word Detection & Keyword Spotting. <https://picovoice.ai/platform/porcupine/>.
- [62] Prebid.org Inc. 2022. Prebid. <https://prebid.org/>.
- [63] Raspberry Pi. 2021. Setting up a Bridged Wireless Access Point. <https://github.com/raspberrypi/documentation/blob/develop/documentation/asciidoc/computers/configuration/access-point-bridged.adoc>.
- [64] Jingjing Ren, Daniel J. Dubois, David Choffnes, Anna Maria Mandalari, Roman Kolcun, and Hamed Haddadi. 2019. Information Exposure From Consumer IoT Devices: A Multidimensional, Network-Informed Measurement Approach. In *Proceedings of the Internet Measurement Conference (IMC '19)*. Association for Computing Machinery, Amsterdam, 267–279.
- [65] Rhasspy. 2022. Rhasspy Voice Assistant. <https://rhasspy.readthedocs.io/>.
- [66] Said Jawad Saidi, Anna Maria Mandalari, Roman Kolcun, Hamed Haddadi, Daniel J. Dubois, David Choffnes, Georgios Smaragdakis, and Anja Feldmann. 2020. A Haystack Full of Needles: Scalable Detection of IoT Devices in the Wild. In *Proceedings of the ACM Internet Measurement Conference (IMC '20)*. Association for Computing Machinery, Pittsburgh, 87–100.
- [67] Hamza Shaban. 2018. Amazon Alexa user receives 1,700 audio recordings of a stranger through 'human error'. <https://www.washingtonpost.com/technology/2018/12/20/amazon-alexa-user-receives-audio-recordings-stranger-through-human-error/>.
- [68] Rita Singh. 2019. *Profiling humans from their voice*. Springer.
- [69] Software Freedom Conservancy. 2022. Selenium. <https://www.selenium.dev/>.
- [70] Sonos, Inc. 2022. Sonos. <https://www.amazon.com/dp/B072NL3N6K>.
- [71] Statista. 2022. Number of households with smart home products and services in use worldwide from 2015 to 2025. <https://www.statista.com/statistics/1252975/smart-home-households-worldwide/>.
- [72] Statista. 2022. Smart speaker devices installed base in the United States from 2017 to 2020. <https://www.statista.com/statistics/794480/us-amazon-echo-google-home-installed-base/>.
- [73] Rahmadi Trimananda, Hieu Le, Hao Cui, Janice Tran Ho, Anastasia Shuba, and Athina Markopoulou. 2022. OVRseen: Auditing Network Traffic and Privacy Policies in Oculus VR. In *31st USENIX Security Symposium (USENIX Security 22)*. USENIX Association, Boston, 3789–3806.
- [74] ttm. 2022. Essential Oil Benefits. <https://www.amazon.com/dp/B074CNX3G8>.
- [75] Janus Varmarken, Hieu Le, Anastasia Shuba, Athina Markopoulou, and Zubair Shafiq. 2020. The TV is Smart and Full of Trackers: Measuring Smart TV Advertising and Tracking. In *Proceedings on Privacy Enhancing Technologies*, Vol. 2020 (2). Sciencio, Virtual, 129–154.
- [76] VCA, INC. 2022. VCA Animal Hospital. <https://amazon.com/dp/B07KYS1Y1X>.
- [77] WaLLy3K. 2022. Pi-hole Blocklist. <https://firebog.net/>.
- [78] Max Willens. 2021. Amid post-cookie confusion, Amazon plans to launch an identifier of its own. <https://digiday.com/marketing/amid-post-cookie-confusion-amazon-explores-launching-an-identifier-of-its-own/amp/>.
- [79] Xeline Development. 2022. Makeup of the Day. <https://amazon.com/dp/B072N6BNB1>.
- [80] Jeffrey Young, Song Liao, Long Cheng, Hongxin Hu, and Huixing Deng. 2022. SkillDetective: Automated Policy-Violation Detection of Voice Assistant Applications in the Wild. In *31st USENIX Security Symposium (USENIX Security 22)*. USENIX Association, Boston, 1113–1130.
- [81] YouVersion. 2022. YouVersion Bible. <https://www.amazon.com/dp/B017RXFNKY>.
- [82] Eric Zeng, Tadayoshi Kohno, and Franziska Roesner. 2021. What Makes a "Bad" Ad? User Perceptions of Problematic Online Advertising. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (Yokohama, Japan) (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 361, 24 pages. <https://doi.org/10.1145/3411764.3445459>
- [83] Jiang Zhang, Konstantinos Psounis, Muhammad Haroon, and Zubair Shafiq. 2022. HARPO: Learning to Subvert Online Behavioral Advertising. In *29th Annual Network and Distributed System Security Symposium*. Internet Society, San Diego.

APPENDIX

A High absolute bid values with only skill installation

This Appendix expands on the discussion of ad bid values after skill installation discussed in Section 5.2.2. We note that the absolute bid values for vanilla and interest personas with only skill installation are higher than that of vanilla and interest personas with user interaction. While it is impossible to know the exact reason for high bid values for personas with only skill installation, a possible explanation could be our data collection during the holiday season, i.e., around Christmas 2021.⁷ To rule out the impact of

⁷We collected advertisements (including bids) both before and after interacting with the skills from 12/08/21 to 12/22/21 and 12/28/21 to 02/04/22, respectively. We installed the skills on 12/07/21 and we interacted with the skills on 12/27/21.

Persona	No Interaction	Interaction
Connected Car	0.364	0.311
Dating	0.519	0.297
Fashion & Style	0.572	0.404
Pets & Animals	0.492	0.373
Religion & Spirituality	0.477	0.231
Smart Home	0.452	0.349
Wine & Beverages	0.418	0.522
Health & Fitness	0.564	0.826
Navigation & Trip Planners	0.533	0.268
Vanilla	0.539	0.232

Table 11: Mean bid values without and with interaction across interest and vanilla personas that were collected close to each other.

holiday season, we compare the bids values with only skill installation and with skill interaction that were collected close to each

other. Specifically, we compare the bids from last three iteration of without interaction with bids from first three iterations of with interaction, that were crawled within a close time span.⁸ Table 11 presents mean bid values without and with user interaction. It can be seen that the interest personas with interaction receive higher bids than vanilla persona. Whereas no discernible differences exist for without interaction configurations.

Although the timing affects the bid values, we believe that it does not impact our findings. Specifically, our objective is to measure the effect of treatment, i.e., skill installation or interaction, on interest (treatment) personas as compared to the vanilla (control) persona. The relative comparison of bid values between vanilla (control) and interest (treatment) personas suffices to measure the effect of treatment. It means that, if we see statistically significant differences in bid values between vanilla (control) and interest (treatment) personas, we can confidently attribute the differences to the applied treatment, i.e., skill installation or interaction.

⁸From 12/20/21 to 12/22/21 for skill installation and from 12/28/21 to 01/12/22 for skill interaction.

